

概率框架下多特征显著性检测算法

李维鹏

摘要: 为提升多特征条件下图像的显著性检测精度,以显著图的联合概率分布为基础,结合先验知识,设计一种概率框架下的多特征显著性检测算法。首先分析了单一特征显著性检测的潜在缺陷,继而推导出多特征下显著图的联合概率分布;然后根据显著图的稀有性,稀疏性,紧凑性与中心先验推导出显著图的先验分布,并使用正态分布假设简化了显著图的条件分布;随后根据显著图的联合概率分布得到其极大后验估计,并基于多阈值假设构建了分布参数的有监督学习模型。数据集实验表明:相比于精度最高的单一特征显著性检测方法,多特征算法在有监督和启发式方法下的平均误差降低了6.98%和6.81%,平均F-measure提高了1.19%和1.16%;单幅图像的多特征融合耗时仅为11.8ms。算法精度较高,实时性好,且可根据不同任务选择所需的特征类别与先验信息,能够满足多特征显著性检测的性能要求。

关键词: 显著性检测; 联合概率分布; 多特征融合; 极大后验估计

Saliency Detection with Multi-features in Probability Framework

Abstract: In order to improve the precision of visual saliency detection with multi-features, a multi-feature integration algorithm is proposed based on the joint probability distribution of saliency map and combined with priori knowledge. Firstly, the potential defects of single feature saliency detection are analyzed, and the joint probability distribution of saliency maps with multiple features is deduced. Secondly, the priori distribution of the saliency map is deduced based on the rarity, sparsity, compactness and center priori of the saliency map, and the condition distribution of the saliency map is simplified based on the assumption of normal distribution. Then the maximum a posteriori estimation is obtained from the joint probability distribution of the saliency map, and a supervised learning model of the distribution parameters is constructed based on the multi-threshold hypothesis. Experiments show that compared to the highest-precision saliency detection method on single feature, the mean average error of the multi-feature algorithm under the supervised and heuristic method is decreased by 6.98% and 6.81%, and the average F-measure is improved by 1.19% and 1.16%. And the multi-feature integration of single image takes only 11.8ms. The algorithm has high accuracy and real-time performance, and can be combined with the required features and different prior information according to the task. It meets the requirements of saliency detection with multi-features.

Key words: saliency detection; joint distribution; multi-feature integration; maximum a posteriori estimation

1 引言

人类感知图像时会有选择性的关注图像中较为显著的区域,并以显著区域为基础添加细节信息,从而完成高级信息的提取^[1]。关于该现象的研究涵盖了生物学、脑科学和计算机视觉等众多领域^[2]。作为信号处理和机器视觉的一项基本问题,显著性检测从经典视觉注意力模型到现在已历经了数十年的研究,并广泛地用于注视点预测、目标检测、场景分类、内容重要性评价等方面。

当前大多数的显著性检测模型主要使用底层视觉特征结合先验知识对显著性进行评估^[10],底层视觉特征如颜色、纹理、粗大轮廓等^[5],先验知识包

含稀有性、稀疏性、紧凑性、中心先验等^[1]。早期算法大多使用单一或为数不多的特征进行显著性检测,诸如 R. Achanta 提出的局部对比度法^[6]、X. Hou 提出的频谱残差法^[4]、J. Harel 等提出的基于图方法的显著性检测^[7]等。而随着新的特征不断被提出,显著性检测迈向了多特征化的道路^{[11]-[14]},尤其是 CNN 的出现^[5],更是突破了传统手工设计特征的局限,转而从图像中自主学习特征,但 CNN 模型受到有监督学习所需巨量样本的限制,难以用于样本较少的情形和启发式显著性检测。

为了将多特征高效的整合并用于显著性检测任务中,本文从计算显著图的全概率公式着手,结合正态分布假设,提出一种融合多特征的显著性检测

架构,如图1所示。该架构既可以用于有监督显著性检测,也可以用于无监督的启发式显著性检测。且融合算法计算量较小,易于并行化和拓展特征,具有较高的实用意义。

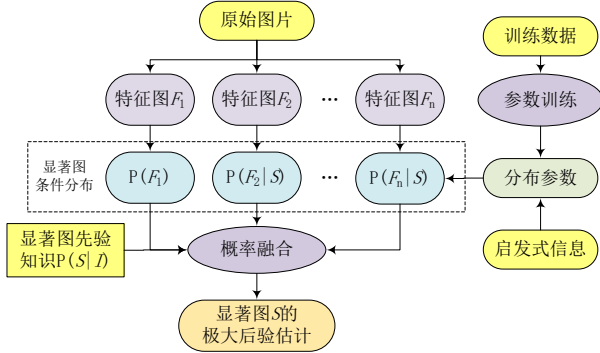


图1 多特征显著性检测算法架构
Fig.1 Algorithm architecture of multi-feature saliency detection

2 显著图的概率表示

2.1 单一特征显著图的概率表示

在显著性检测方面,文献通常会给出由原始图像到中间特征的计算方法,以及由特征到显著图的计算流程,以便清晰的展现算法思想。有些算法也会直接将特征图作为显著图输出。为便于分析对比,首先将这些文献的显著图的计算从形式上归纳为极大后验估计:

$$S = \arg \max P(S|I) \quad (1)$$

其中 I 表示原始图像, S 表示对应的显著图。

在概率框架下,各个算法的区别仅在于概率分布 $P(S|I)$ 的不同。进一步地,对于某种特征,可由原始图像 I 得到对应的特征强度图 F ,显著图的概率分布可表示为:

$$P(S|F, I) = \frac{P(F|S)P(S|I)}{P(F|I)} \quad (2)$$

其中 $P(F|S)$ 为显著图 S 关于特征图 F 的似然概率; $P(S|I)$ 为显著图关于原图 I 的条件概率,本文将其作为 S 的先验分布; $P(F|I)$ 为特征图 F 关于原图 I 的概率,由于 F, I 均为已知量,故本文将该项作为常数。

式(2)可以看出,给定由原始图像 I 得到特征图 F 的方式后,显著图 S 由特征强度图 F 与先验信息决定。且式(2)表明当使用单一特征时,显著图关于特征图的似然概率 $P(F|S)$ 的离散程度——尤其是方差——对显著性检测结果的影响是单向的。若 $P(F|S)$ 最大时 S 的部分元素与真值的偏差远大于给定方差时, S 的误差将难以修正,故单一特征存在着以偏概全的风险。因此,我们认为应当使用多类特征综合估计显著图,同时权衡各个特征在显著性检测任务中对显著性度量的联合分布,以实现更为精准的显著性检测。

2.2 多特征显著图的概率表示

首先考虑多特征下,显著图 S 的概率问题。所选的特征类型为 $\mathcal{F} = \{f_1, \dots, f_n\}$, 对应各个特征的强度图为 $\{F_1, \dots, F_n\}$, 显著性检测问题表示为:

$$S = \arg \max P(S|F_1, \dots, F_n, I) \quad (3)$$

展开式(3),得到:

$$P(S|F_1, \dots, F_n, I) \propto P(S|I) \prod_{i=1}^n P(F_i|S) \quad (4)$$

其中 $P(S|I)$ 表示显著图先验分布,可包含显著图的稀有性、稀疏性、紧凑性、中心先验等信息; $P(F_i|S)$ 特征图 F 关于显著图 S 的条件分布,用于表示显著图对特征图的相关程度。

尤其需要注意的是,当先验分布不为均匀分布时,存在有效的显著图先验信息,此时在显著性检测中需要考虑与先验分布相一致的先验信息才能保证算法的准确性。

式(4)中各个先验分布和条件分布都有着各自的具体含义,下面逐一进行说明。

2.2.1 显著图先验分布

根据需要,显著图先验分布可有选择性地表示稀有性、稀疏性、紧凑性、中心先验等信息。假设上述性质的分布之间相互独立,可得:

$$P(S|I) = \prod_{v=1}^m P_v(S|I) \quad (5)$$

其中 $P_v(S|I)$ 表示第 v 种先验信息。

显著图的每种先验信息的表示方法都各有区别,常见的先验信息包括:

1)、稀有性 表示显著图的能量较小,该性质

可由显著图的 F 范数 $\|S\|_F$ 表示, $\|S\|_F$ 越小显著图稀有性越强;

2)、稀疏性 表示显著区域的面积较小, 可由显著图的 L0 范数 $\|S\|_{L0}$ 表示, 由于 L0 范数难以优化, 该性质可凸近似地表示为 $\|\log(S+1)\|_{L1}$, 范数越小显著图稀疏性越强;

3)、紧凑性 表示显著区域较为集中, 可由显著图的 F 范数与其高斯滤波后的 F 范数平方之差 $\|S\|_F^2 - \|S_G\|_F^2$ 表示, 该值越小显著图紧凑性越强;

4)、中心先验 反映了人类视觉对视场中心区域较为关注的性质。空间域中, 该性质可由显著图 S 与中心化遮罩 Γ 的元素积 (Hadamard product):

$$S_c = S \circ \Gamma \quad (6)$$

表示, 其中 S_c 表示添加中心先验后的显著图。

需要说明的是, 由于上述稀有性、稀疏性和紧凑性选用范数作为评价指标, 该指标的优化具有单峰特性; 而正态分布与负指数分布是单峰分布中计算相对简单的形式。综合考虑模型精度和计算复杂度, 上述 3 种先验信息的分布使用正态分布或负指数分布模型:

$$\begin{cases} \|S\|_F \sim \mathcal{N}(0, \sigma_r^2) \\ \|\log(S+1)\|_{L1} \sim \text{Exp}(-\beta_s) \\ \|S\|_F^2 - \|S_G\|_F^2 \sim \text{Exp}\left(-\frac{\alpha_c}{2}\right) \end{cases} \quad (7)$$

对于更加复杂的先验分布——指数分布族的分析, 将在第 3 章中予以讨论。

2.2.2 显著图条件分布

条件分布即显著图 S 对给定类型特征 f_i 的特征图 F_i 的条件分布:

$$P(F_i | S) \quad (8)$$

该分布反映了显著图 S 与特征 f_i 的相关程度。例如当 $F_i | S$ 满足正态分布 $F_i | S \sim \mathcal{N}\{F_i(S), D_i\}$ 时, 方差 D_i 越大, S 与特征图 F_i 的相关程度也就越低。

式(4)中的多个条件分布对应于不同类型的特征, 该表示下通过选取恰当的分布参数能够有机融合各个特征的优点, 并规避特征的不足, 从而达到取长补短的目的。

3 显著图的极大后验估计

3.1 指数分布族下的显著性检测

指数分布族是一类常见分布, 涵盖了正态分布、二项分布、泊松分布、伽马分布等诸多分布函数。为便于描述各个分布为指数分布族时, 显著图的极大后验估计, 首先将显著图 S 展开为向量形式 s , 则单个类型的先验分布描述为:

$$P_v(s | \boldsymbol{\eta}_v) = \exp\{\boldsymbol{\eta}_v^T \mathbf{t}_v(s) + b_v(s) - \rho_v(\boldsymbol{\eta}_v)\} \quad (9)$$

其中 $\boldsymbol{\eta}_v$ 为分布参数, $\mathbf{t}_v(s)$ 为 s 的向量函数, $b_v(s)$ 为 s 的标量函数, $\rho_v(\boldsymbol{\eta}_v)$ 为 $\boldsymbol{\eta}_v$ 的标量函数。

将式(4)给出的显著图分布函数对数化。多特征显著性检测可表示为:

$$S = \arg \max \left[\log P(S | \mathbf{I}) + \sum_{i=1}^n \log P(F_i | S) \right] \quad (10)$$

解得:

$$\begin{aligned} \sum_{i=1}^n \frac{1}{P(S | \mathbf{I})} \frac{\partial P(S | \mathbf{I})}{\partial S} + \\ \frac{1}{P(F_i | S)} \frac{\partial P(F_i | S)}{\partial S} = 0 \end{aligned} \quad (11)$$

将(9)代入(11)得到:

$$\begin{aligned} \sum_{i=1}^n \frac{\partial}{\partial S} \varphi_i(s, \hat{s}_i; \boldsymbol{\theta}_i) + \\ \sum_{v=1}^m \left[\frac{\partial \mathbf{t}_v^T(s)}{\partial S} \boldsymbol{\eta}_v + \frac{\partial b_v(s)}{\partial S} \right] = 0 \end{aligned} \quad (12)$$

其中 $\varphi_i(s, \hat{s}_i; \boldsymbol{\theta}_i)$ 为显著图条件分布的负对数形式, \hat{s}_i 为特征 f_i 输出的显著图 (向量形式), $\boldsymbol{\theta}_i$ 为 \hat{s}_i 的分布参数。

下面针对指数分布族的各种情况予以讨论。

3.2 正态分布与负指数分布下的解析求解

作为指数分布族中的特例, 正态分布和负指数分布的极大后验估计具有简洁的解析解。将式(7)代入(12), 频域下, 显著图 S 的极大后验估计满足:

$$\begin{aligned} \sum_{i=1}^n \mathbf{W}_i \circ (S - \hat{S}_i \circ \boldsymbol{\Gamma}_i) + \\ [\sigma_r^{-2} S + \alpha_c (S - S \circ \mathbf{G}^{[2]}) + \beta_s (S + 1)^{[-1]}] = 0 \end{aligned} \quad (13)$$

其中 \mathbf{W}_i 为第 i 个特征的权重矩阵, $\boldsymbol{\Gamma}_i$ 为线性系数, $\boldsymbol{\theta}_i = \{\mathbf{W}_i, \boldsymbol{\Gamma}_i\}$, \mathbf{G} 为高斯低通滤波器, $\mathbf{G}^{[p]}$ 表示 \mathbf{G} 所有元素的 p 次幂构成的矩阵 (或向量), 其余具有相同形式的上标同理; 中括号中的各项表示显著图的

先验信息， σ_r 、 α_c 、 β_s 分别表示显著图稀有性标准差、中心化参数、稀疏性参数。

式(13)为 S 的 2 阶函数，容易求得解析解。当先验分布不为正态分布或负指数分布时，使用 Newton 法或 BFGS 等算法也可快速求得数值解。

3.3 指数分布族下的 Newton 数值求解

上一节使用正态分布与负指数分布作为先验信息的概率分布，并基于该分布假设得到多特征显著图的解析解。然而该假设具有一定的局限性，若考虑到更加一般化的指数分布族，显著图的解析解将难以求取，此时可以借助 Newton 法获得数值解。

为便于描述 Newton 法求解过程，首先将显著图 S 展开为向量形式 s 。设第 k 次迭代结果为 s_k ，由式(12)得到其 Newton 迭代满足：

$$s_{k+1} = s_k - H_k^{-1} \varepsilon(s_k) \quad (14)$$

其中：

$$\varepsilon(s_k) = \sum_{i=1}^n \frac{\partial}{\partial s} \varphi_i(s, \hat{s}_i; \theta_i) + \sum_{v=1}^m \left[\frac{\partial t_v^T(s)}{\partial s} \eta_v + \frac{\partial b_v(s)}{\partial s} \right] \quad (15)$$

H_k 为 Hessian 矩阵，满足：

$$H_k = \frac{\partial \varepsilon(s_k)}{\partial s^T} = \sum_{i=1}^n \frac{\partial^2}{\partial s \partial s^T} \varphi_i(s, \hat{s}_i; \theta_i) \bigg|_{s=s_k} + \sum_{v=1}^m \left[\frac{\partial^2 t_v^T(s) \eta_v}{\partial s \partial s^T} + \frac{\partial^2 b_v(s)}{\partial s \partial s^T} \right] \bigg|_{s=s_k} \quad (16)$$

一般来讲 Newton 法的计算量集中于 Hessian 矩阵的求解。在式(16)所示 Hessian 矩阵中， $\frac{\partial^2 t_v^T(s) \eta_v}{\partial s \partial s^T}$ 的求解将会是最大的问题：若 $t_v(s)$ 中 s 的各个元素耦合不严重，其二阶偏导容易求解；但当 $t_v(s)$ 中的元素耦合严重时，该项 2 阶偏导数的解算将十分复杂，此时 Newton 法不再适用，需要考虑更加高效的 BFGS 算法。

3.4 指数分布族下的 BFGS 数值求解

BFGS 是一种高效的拟牛顿法，广泛地用于 Hessian 矩阵难以求解的情形。其主要思想是用相邻两次迭代的 s 与 $\varepsilon(s)$ 变化估计 Hessian 矩阵的近似值，以避免求解 Hessian 矩阵的复杂计算。

设 $\Delta s_k = s_{k+1} - s_k$ ， $\Delta \varepsilon_k = \varepsilon(s_{k+1}) - \varepsilon(s_k)$ ，Hessian 矩阵的迭代更新表示为：

$$B_{k+1}^{-1} = \left(E - \frac{\Delta s_k \Delta \varepsilon_k^T}{\Delta \varepsilon_k^T \Delta s_k} \right) B_k^{-1} \left(E - \frac{\Delta \varepsilon_k \Delta s_k^T}{\Delta \varepsilon_k^T \Delta s_k} \right) + \frac{\Delta s_k \Delta s_k^T}{\Delta \varepsilon_k^T \Delta s_k} \quad (17)$$

其中 E 为单位对角阵。

s 的更新为：

$$s_{k+1} = s_k - B_k^{-1} \varepsilon(s_k) \quad (18)$$

Algorithm 1: 显著图的 BFGS 迭代求解

Input: 各个特征输出的显著图 \hat{s}_i ，条件分布系数

θ_i ；先验信息分布系数 η_v

Initialize: $k = 0$ ， s_0 ， $\varepsilon(s_0)$ ， $B_0 = E$

- 由式(18)更新 s_{k+1}
- 由式(15)更新 $\varepsilon(s_{k+1})$
- 判断 $\varepsilon(s_{k+1})$ 是否达到停止条件，若达到条件跳出
- $\Delta s_k = s_{k+1} - s_k$ ， $\Delta \varepsilon_k = \varepsilon(s_{k+1}) - \varepsilon(s_k)$
- 由式(17)更新 B_k
- $k := k + 1$

Output: 显著图估计结果 s

图 2 是显著图的 Newton 数值求解与 BFGS 数值求解的收敛过程对比。其中稀有性和稀疏性选用 Dirichlet 分布，紧凑性选用 Beta 分布，输入图像的尺寸归一化为 256×256 。从结果可以明显看出，在相同的时间内，BFGS 在初值不理想的情况下，收敛速度高于 Newton 法，展现了 BFGS 在简化 Hessian 矩阵计算方面的有效性。

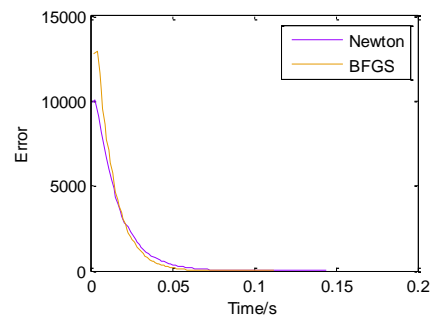


图 2 Newton 法与 BFGS 显著图优化效率对比

Fig.2 Efficiency comparison of Newton algorithm and BFGS in saliency map optimization

4 基于有监督学习的显著性检测

4.1 指数分布族下的参数训练

融合多特征的显著性检测不仅能够作为启发式

算法在无监督情况下使用，也可结合已知的前景标注结果进行有监督学习，优化显著性检测精度。将式(12)中的各个参数作为未知量，根据最小二乘准则进行参数训练：

$$\min_{(\eta, \theta)} J = \sum_{k=1}^l \|s_k(\widehat{s}_{i,k}; \eta, \theta) - \bar{s}_k\|_F^2 \quad (19)$$

其中 l 为训练集图像数目，下标 k 表示训练数据中的第 k 幅图像，其显著图标注值为 \bar{s}_k ，估计值为 s_k ， $\eta = \{\eta_v, v=1, \dots, m\}$ 表示待求解先验分布参数的集合， $\theta = \{\theta_i, i=1, \dots, n\}$ 为待求解的第 i 类特征融合参数的集合。

对式(19)求关于 η, θ 的偏导数，解得：

$$\begin{cases} \frac{\partial J}{\partial \theta_i} = \sum_{k=1}^l \frac{\partial s_k}{\partial \theta_i} (s_k - \bar{s}_k), i=1, 2, \dots, n \\ \frac{\partial J}{\partial \eta_v} = \sum_{k=1}^l \frac{\partial s_k}{\partial \eta_v} (s_k - \bar{s}_k), v=1, 2, \dots, m \end{cases} \quad (20)$$

由于 s_k 能使式(15)误差函数 $\varepsilon(s_k)$ 为 0，则 $\frac{\partial \varepsilon}{\partial \theta_i}$ ， $\frac{\partial \varepsilon}{\partial \eta_v}$ 满足：

$$\begin{cases} \frac{\partial \varepsilon}{\partial \theta_i} = -\frac{\partial \varepsilon}{\partial \theta_i} \left[\frac{\partial \varepsilon}{\partial s_k} \right]^{-1}, i=1, 2, \dots, n \\ \frac{\partial \varepsilon}{\partial \eta_v} = -\frac{\partial \varepsilon}{\partial \eta_v} \left[\frac{\partial \varepsilon}{\partial s_k} \right]^{-1}, v=1, 2, \dots, m \end{cases} \quad (21)$$

其中：

$$\begin{cases} \frac{\partial \varepsilon}{\partial s_k} = \sum_{i=1}^n \frac{\partial^2 \varphi_i(s_k, \widehat{s}_{i,k}; \theta_i)}{\partial s_k \partial s_k^T} + \sum_{v=1}^m \left[\frac{\partial^2 t_v(s_k) \eta_v}{\partial s_k \partial s_k^T} + \frac{\partial^2 b_v(s_k)}{\partial s_k \partial s_k^T} \right] \\ \frac{\partial \varepsilon}{\partial \theta_i} = \frac{\partial \varphi_i(s_k, \widehat{s}_{i,k}; \theta_i)}{\partial \theta_i \partial s_k^T}, i=1, 2, \dots, n \\ \frac{\partial \varepsilon}{\partial \eta_v} = \frac{\partial t_v(s_k)}{\partial s_k^T}, v=1, 2, \dots, m \end{cases} \quad (22)$$

当式(19)所示偏导数等于 0 的方程容易求解时，可求得 η, θ 解析解或 BFGS 迭代求得数值解；若其形式非常复杂，可采用批量梯度下降法求得数值解。

4.2 正态分布与负指数分布下的参数训练

当显著图的条件满足正态分布，各项先验分布满足式(7)所示正态分布或负指数分布时，可由(20)

解得：

$$\begin{cases} \sum_{i=1}^n \sum_{k=1}^l \mathbf{W}_i \circ \mathbf{F}_i \circ \widehat{s}_{i,k} = \sum_{i=1}^n \sum_{k=1}^l s_k \circ \widehat{s}_{i,k} \\ \mathbf{W}_i = \frac{l}{\sum_{k=1}^l |s_k - \widehat{s}_{i,k} \circ \mathbf{F}_i|^2}, i=1, 2, \dots, n \end{cases} \quad (23)$$

其中 $|\bullet|^2$ 表示各个元素绝对值的平方。

式(23)关于 $\mathbf{W}_i \circ \mathbf{F}_i$ 的式子可看成线性方程组进行求解。线性系数 \mathbf{F}_i 可根据 $\mathbf{W}_i \circ \mathbf{F}_i$ 的值以及式(23)关于 \mathbf{W}_i 的式子联合求解。

先验参数满足：

$$\begin{cases} \sigma_r^2 = \frac{\sum_{k=1}^l \|s_k\|_F^2}{l \dim(s)} \\ \beta_s = \frac{l \dim(s)}{\sum_{k=1}^l \|\log(s_k + 1)\|_{L1}} \\ \alpha_c = \frac{l \dim(s)}{\sum_{k=1}^l (\|s_k\|_F^2 - \|s_k \circ \mathbf{G}\|_F^2)} \end{cases} \quad (24)$$

其中 $\dim(s)$ 为 s 的维度。

5 实验与结果

为验证概率框架下多特征融合显著性检测的有效性，我们挑选了 6 种显著性检测算法参与多特征融合及对比实验，实验所用数据集为 ECSSD^[12]，共包含 1000 幅图像。本文实验的 matlab 代码已发布于 github。

5.1 待融合显著性特征性能分析

特征融合实验所用的显著性特征及方法包括：Yan Q 等人^[11]的层次显著性检测（记为 HS）以及改进版^[12]（记为 CHS）；Cheng M M 等人^[13]的局部直方图对比法（记为 RC）；Jiang H 等人^[14]的局部特征综合法（记为 DRFI）。4 种显著性检测算法的示例效果如图 3 所示，其中左上角为原图，后续依次是：HS，CHS，RC，DRFI 的结果，最后一幅图为人标注（GT）。

在此我们仅对 HS、RC 这 2 种方法进行介绍，目的在于展现不同方法的优缺点，为多特征融合显

著性检测的启发式参数选取提供参考依据。



图3 4种原始算法的显著性检测结果

Fig.3 Saliency maps obtained from 4 original detection algorithms

5.1.1 层次显著性检测

层次显著性检测(HS)为典型的多层综合处理方法，其计算多个尺度的局部对比度，随后使用最小化能量函数的形式进行综合。

首先在不同尺度下对图像进行分割，随后计算各层的局部对比度：

$$C_i = \sum_{j=1}^n w(R_j) \phi(i, j) \|c_i - c_j\|_2 \quad (25)$$

其中 c_i 与 c_j 表示分割区域 R_i 与 R_j 的颜色， $w(R_j)$ 为 R_j 的像素面积。更大的面积有着更大的局部对比度。 $\phi(i, j)$ 为区域 R_i 与 R_j 之间的距离权重，表示距离越近相互影响越大。

随后将对比度强度图进行中心化：

$$\bar{s}_i = C_i \cdot H_i \quad (26)$$

在获得各层对比度强度图后，根据最小化以下能量函数得到最终的显著图：

$$E(S) = \sum_l \sum_i E_D(s_i^l) + \sum_l \sum_{i, R_i^l \subseteq R_j^{l+1}} E_S(s_i^l, s_j^{l+1}) \quad (27)$$

其中 $E_D(s_i^l)$ 为显著性置信度， $E_S(s_i^l, s_j^{l+1})$ 为层次能量一致性。两者分别满足：

$$E_D(s_i^l) = \beta^l \|s_i^l - \bar{s}_i^l\|_2^2 \quad (28)$$

$$E_S(s_i^l, s_j^{l+1}) = \lambda^l \|s_i^l - s_j^{l+1}\|_2^2 \quad (29)$$

由于对比度计算时面积越大的区域能量越大，该方法对尺度较大的显著区域更加敏感，而对小显著目标检测的置信度较低。

5.1.2 局部特征综合法

该方法使用 Lab 色彩空间。定义像素的对比度为该像素与图像其余所有像素的颜色差异之和：

$$S(r_k) = \sum_{r_i \neq r_k} \exp(-D_s(r_k, r_i) / \sigma_s^2) w(r_i) D_r(r_k, r_i) \quad (30)$$

其中 $D_s(r_k, r_i)$ 为区域 r_k, r_i 之间的空间距离。 $w(r_i)$ 为区域 r 权重， $D_r(r_k, r_i)$ 表示 r_k, r_i 之间的颜色差异，定义为：

$$D_r(r_1, r_2) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} f(c_{1,i}) f(c_{2,j}) D(c_{1,i}, c_{2,j}) \quad (31)$$

其中 $f(c_{k,i})$ 是第 i 种颜色在 k 区域 r_k 的概率。

由于采用局部统计进行显著性检测，该方法对于大面积的局部特殊颜色区域非常敏感，但对较小的显著区域缺乏敏感性。

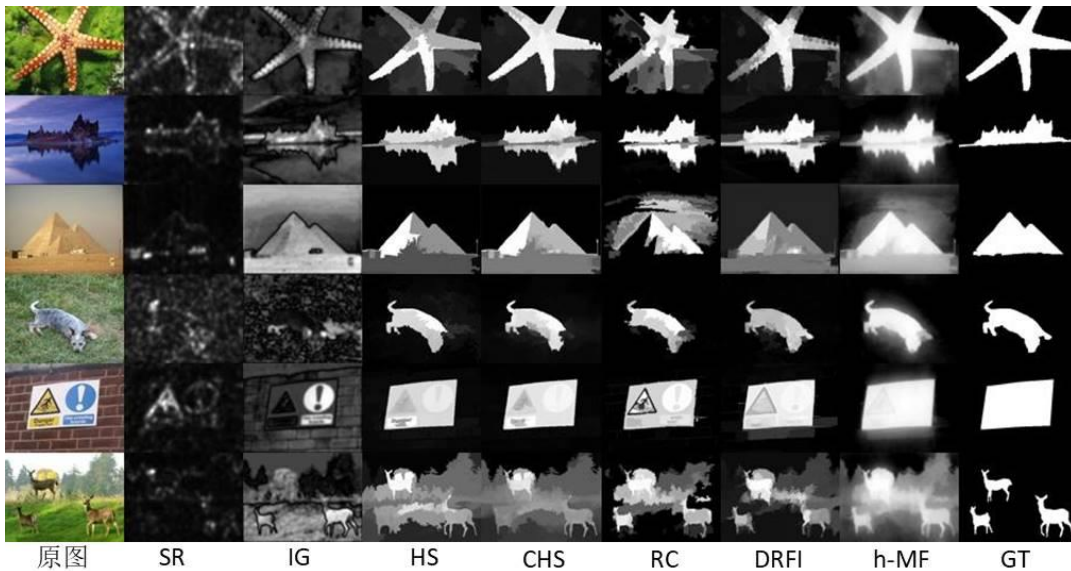


图4 6种对比算法与多特征融合算法的显著图对比

Fig.3 Comparison of saliency maps obtained from 6 contrast algorithms and the MF algorithm

5.2 启发式多特征显著性检测

实验使用 Shi J 的 ECSSD^[12]数据集，特征融合使用：IG，CHS，RC，DRFI 算法。对比算法为：谱残差法（SR），IG，HS，CHS，RC，DRFI。为方便起见，启发式多特征融合算法记为 h-MF。

为了提升显著性检测结果的全局性，选择在频域进行特融合。根据上一节对各个显著性检测算法特性的分析，可针对性地设计各个算法的融合参数。

根据对结果的观察分析，CHS 倾向于敏感大尺度目标，中低频性能良好，且总体置信度 $\|W_i\|_F$ 较高，对应 W_i 低频部分较大，高频部分较小。作为全局统计直方图，RC 具有较好的全局特性，但丢失了小尺度的局部信息，其权重矩阵 W_i 高频较低。DRFI 纳入了多种特征，其全局性能良好， W_i 整体较大。

6 种原始算法与融合算法的显著性检测结果对比如图 4 所示，其中最左边为原图，后续依次是：SR，IG，HS，CHS，RC，DRFI、h-MF 的显著性检测结果，最后一幅图为人工标注（GT）。注意到，h-MF，将各个算法的优点较好地结合在一起。总体上提升了显著性检测效果。

图 5 为 7 种算法在 ECSSD 数据集实验中的准确率-召回率（PR）曲线，其中黑色曲线为 h-MF 的结果，表 1 展示了由 PR 曲线所得的各个算法平均精度（MAP）、平均召回率（MAR）与平均 F-measure（Mean F-m）。从结果中看到，相比于对比算法，h-MF 算法具有显著优势。

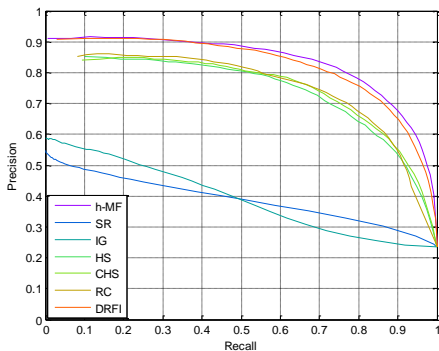


图 5 h-MF 及对比算法的准确率-召回率曲线

Fig.4 Precision-recall of h-MF and algorithms compared

5.3 有监督多特征显著性检测

当拥有训练数据时，可根据 4.1 节或 4.2 节方法对模型参数进行有监督训练。

各个算法在 ECSSD 数据集实验中的准确率-召回率曲线如图 6 所示。相比于其余算法，有监督多特征融合算法具有明显优势。

需要说明的是，Sup-MF 的 MAP 低于 h-MF 的原因在于，有监督方法使用式(19)作为目标函数，而该函数表示的是预测值与标注值的均方误差，并非对 MAP 直接进行优化（因为对 MAP 进行优化需要预先估计各个阈值对应的召回率，而这种预估通常是不准确的）。由于训练目标与评价标准的这一差异，有监督方法的 MAP 低于启发式方法。

表 2 展示了有监督多特征融合（记为 sup-MF）算法的训练耗时和测试耗时（不包括特征检测用时），训练和测试均使用 1000 幅图像。训练过程总耗时为 42.555s；1000 幅图像的测试过程总耗时为 11.838s，每幅图像平均测试耗时仅有 0.0118s，且标准差较小，性能稳定。

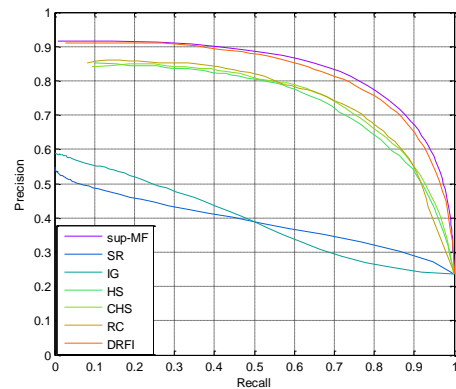


图 6 sup-MF 及对比算法的准确率-召回率曲线

Fig.4 Precision-recall of sup-MF and algorithms compared

表 1 各个算法的 MAP, MAR 与平均 F-measure

Tab.1 MAP, MAR and mean F-measure of each algorithm

| 算法 | MAP | MAR | Mean F-m |
|-----|--------|--------|----------|
| SR | 0.3882 | 0.7088 | 0.5017 |
| IG | 0.3930 | 0.6679 | 0.4949 |
| HS | 0.7423 | 0.8714 | 0.8017 |
| CHS | 0.7503 | 0.8834 | 0.8114 |

| | | | |
|--------|---------------|---------------|---------------|
| RC | 0.7537 | 0.8753 | 0.8099 |
| DRFI | 0.8238 | 0.9048 | 0.8624 |
| h-MF | 0.8361 | 0.9126 | 0.8727 |
| sup-MF | 0.8358 | 0.9124 | 0.8724 |

表 2 1000 幅图像中 MF 算法的训练和测试耗时（不包括特征检测时间）

Tab.2 Time-used by training and test of MF algorithm in 1000 image (excluding feature detection time)

| | 平均耗时(s) | 标准差(s) | 总耗时(s) |
|-------|---------|--------|--------|
| 有监督训练 | 0.043 | 0.012 | 42.555 |
| 测试 | 0.0118 | 0.003 | 11.838 |

6 结论

本文为提升多特征条件下图像的显著性检测精度，以多特征下显著图的联合概率分布为基础，结合先验知识，设计一种概率框架下的多特征融合显著性检测算法。该算法既可使用有监督的显著性检测，也可以用于启发式显著性检测，且可以根据需要方便地扩充特征类别与先验信息，能够满足不同情形下多特征融合显著性检测的性能要求。

为了将特征进行最佳融合，本文推导了多特征下显著图的联合概率分布，以度量各个特征对显著图影响的不确定性，得到更加准确的融合结果。实验结果表明，融合算法在有监督和启发式方法下的平均准确率分别为 0.8361 和 0.8358，相比于精度最高的 DRFI 显著性检测方法误差降低了 6.98% 和 6.81%，平均 F-measure 分别提高了 1.19% 和 1.16 %。

本文使用正态分布假设的目的在于降低显著图极大后验估计的计算量，尽管融合结果从实验的角度而言较为理想，然而该假设显然并非最优，实际情形会更加复杂。我们会在以后的工作中更深入地研究该问题。

参考文献:

- [1] BORHI A, CHENG M M, HOU Q, et al. Salient Object Detection: A Survey[J]. Eprint Arxiv, 2014, 16(7): 3118-3137.
- [2] BORHI A, SIHITE D N, ITTI L. Salient Object Detection: A Benchmark[M]// Computer Vision – ECCV 2012. Springer: Berlin Heidelberg, 2012: 414-429.
- [3] GUO C, MA Q, ZHANG L. Spatio-temporal Saliency detection using phase spectrum of quaternion fourier transform[C]// Computer Vision and Pattern Recognition, 2008. IEEE Conference on. IEEE, 2008: 1-8.
- [4] HOU X, ZHANG L. Saliency Detection: A Spectral Residual Approach[C]// Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on. IEEE, 2007: 1-8.
- [5] 郑云飞, 张雄伟, 曹铁勇, 等. 基于全卷积网络的语义显著性区域检测方法研究[J]. 电子学报, 2017, 45(11): 2593-2601.
ZHEN Y F, ZHANG X W, XIA T Y, et al. The Semantic Salient Region Detection Algorithm Based on the Fully Convolutional Networks[J]. Acta Electronica Sinica, 2017, 45(11): 2593-2601.
- [6] ACHANTA R, ESTRADA F, WILS P, et al. Salient Region Detection and Segmentation[C]// International Conference on Computer Vision Systems. Springer: Berlin, Heidelberg, 2008: 66-75.
- [7] Harel J, Koch C, Perona P. Graph-Based Visual Saliency[C]// International Conference on Neural Information Processing Systems. MIT Press, 2006: 545-552.
- [8] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[M]. IEEE Computer Society, 1998.
- [9] MA Y F, ZHANG H J. Contrast-based image attention analysis by using fuzzy growing[C]// Eleventh ACM International Conference on Multimedia. ACM, 2003: 374-381.
- [10] ACHANDA R, HEMAMI S, Estrada F, et al. Frequency-tuned salient region detection[C]// Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009: 1597-1604.
- [11] Yan Q, Xu L, Shi J, et al. Hierarchical Saliency Detection[C]// Computer Vision and Pattern Recognition. CVPR 2013. IEEE Conference on. IEEE, 2013, 38: 1155-1162.
- [12] Shi J, Yan Q, Xu L, et al. Hierarchical Saliency Detection on Extended CSSD[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 38(4):717.
- [13] Cheng M M, Mitra N J, Huang X, et al. Global Contrast Based Salient Region Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3):569-582.
- [14] Jiang H, Yuan Z, Cheng M M, et al. Salient Object Detection: A Discriminative Regional Feature Integration Approach[J]. International Journal of Computer Vision, 2017, 123(2):251-268.